Effectiveness of Information Disclosure

RAHUL TELANG
CARNEGIE MELLON UNIVERSITY

Carnegie Mellon
Information Security
Policy & Management
Heinz School

Motivation

- A huge problem in information security and privacy arena is incompatible incentives of different parties.
  - Firms have higher incentives to use a user's personal incentives than the user wants.
  - Firms have less incentives to keep users data safe than what is socially desirable.
  - Software vendors have fewer incentives to improve security and release timely patches.
- This creates a significant policy challenge. What can policy makers do to improve market outcomes?

Motivation

- All policy interventions have a downside.
  - Tax or penalize the firm.
  - Hold them liable to losses they cause.
  - Force them to adopt certain standards which are audited and monitored.

- Effectiveness of each of these can be debated.

- Another policy tool commonly used is to force firms (or a third party) to disclose (possibly unpleasant) information. For example,
  - Pollution volumes of a factory.
  - Disclose information about restaurant hygiene.
  - Disclose information about a data breach.
  - Disclose information about a software flaw.

Why Disclosure?

- Disclosure can be considered as “soft paternalism”.
  - Firms are not necessarily taxed or penalized directly. All they are asked is to disclose some information.
- In a competitive market, disclosing information informs customer s about firms practices, which in long run should force firms to do the “right thing”.
- In fact, in competitive markets, sometimes firms voluntarily disclose some information.
  - Though voluntary disclosure does not always happen. Therefore, we need some regulations.
- That said, firms often complain about effectiveness of disclosure policies?
  - They tend to be costly to firms.

Where are we going?

- I will examine two distinct arenas where “disclosure” has been a challenging policy issue.
  - Disclosing information about software vulnerabilities.
  - Disclosing information about data breaches.

Vulnerability Disclosure

- A unique feature of software is that anyone can find flaws (randomly or by exerting effort).

- What options the discoverer has?
  - Inform the vendor and hope the vendor releases a timely patch.
  - Keep quiet.
  - Make the information public.
  - Sell this information.

- Many discoverers do not trust the vendors to provide timely and reliable patches. In many instances, they resort to disclosing vulnerability information in public forums in the hope of pressuring vendors.

- But making this information public causes potential harm to every user of that software (again a unique feature of software)

Disclosure Debate

- This has created some controversies regarding what is the right thing to do
  - Full disclosure
  - Partial disclosure
  - Involving a middleman to manage disclosures (for example, CERT).
- Eventually firms like TippingPoint and iDefence started buying and selling vulnerability information.

What's the tension?

- Security vulnerabilities are costly to vendors and customers. After all vulnerability is an unpleasant information (a "bad" good).
    - Vendors incur cost of patching. One would expect that more time they have for patching less it costs. In short, they would like to delay the patch as much as possible.
    - Vendors' customer incur loss if the vulnerabilities are exploited when they are breached. Vendor cares about customer loss (reputation loss, sometimes contractual obligations). If they cared about customer losses a lot (completely) then the solution is simple. Just inform the vendor and it will do the right thing.
- What does public disclosure do?
    - Increase customer costs. Disclosure leads to more attacks.
    - Increase vendor cost. After all vendor cares about customer costs (atleast somewhat).
- Therefore, many studies and industry practices suggest users should keep quite for some time and allow vendors to release patch, and then threaten disclosure.
    - For example, CERT keeps quite for 45 days.

Theory is nice but….

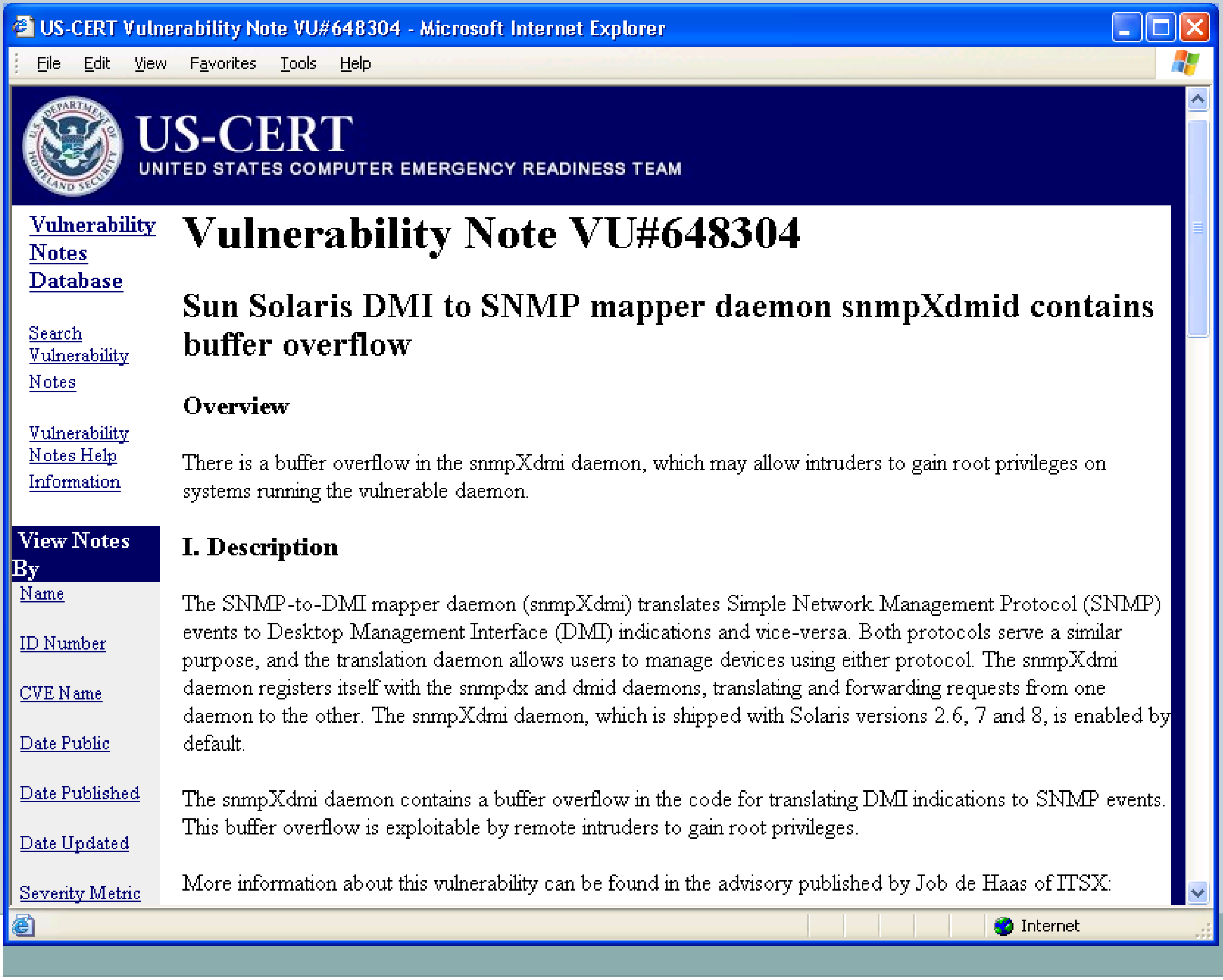
- As is common in any policy formulation, setting up a theoretical model is useful and informative.
  - But there are too many moving parts in the environment that cannot be readily modeled.
  - So the deeper insights into the effectiveness of a policy depends on empirical and experimental work.

- Some of my work has tried to measure how vendor's decision to release patches is affected by disclosure (Arora, Krishnan, Telang and Yang 2010, *Information Systems Research*, Arora, Telang, Xu 2006, *Management Science*)

Data

- Vulnerabilities are published by many sources. We focus on CERT/CC and Securityfocus (SF).
- CERT/CC researches the vulnerability when a user notifies it, and then contacts the vendor and provide it with “protected period”.
- Securityfocus (an online open forum) has less quality control and does not provide any “protected period” unless individuals choose to do so. In short, individuals can post the vulnerability information openly.
- Vulnerabilities published by SF and/or CERT/CC from 9/26/2000 to 8/11/2003.
  - Vendor Notification date (CERT provided us this data and from SF website)
  - Patch release date (from vendor websites and from CERT and SF).
- Vendor information
  - from Hoover’s online and vendor’s website
- Vulnerability characteristics
  - from the national vulnerability database (NVD, previously ICAT database)
- After cleaning up the data we have 1280 observations, related to 255 unique vendors and 303 unique vulnerabilities.

Example of vulnerability publication by CERT



US-CERT Vulnerability Note VU#648304 - Microsoft Internet Explorer
File
Edit
View
Favorites
Tools
Help
US-CERT
UNITED STATES COMPUTER EMERGENCY READINESS TEAM
Vulnerability Notes Database
Search Vulnerability Notes
Vulnerability Notes Help Information
View Notes By
Name
ID Number
CVE Name
Date Public
Date Published
Date Updated
Severity Metric
Internet

Vulnerability Note VU#648304

Sun Solaris DMI to SNMP mapper daemon snmpXdmid contains buffer overflow

Overview

There is a buffer overflow in the snmpXdmi daemon, which may allow intruders to gain root privileges on systems running the vulnerable daemon.

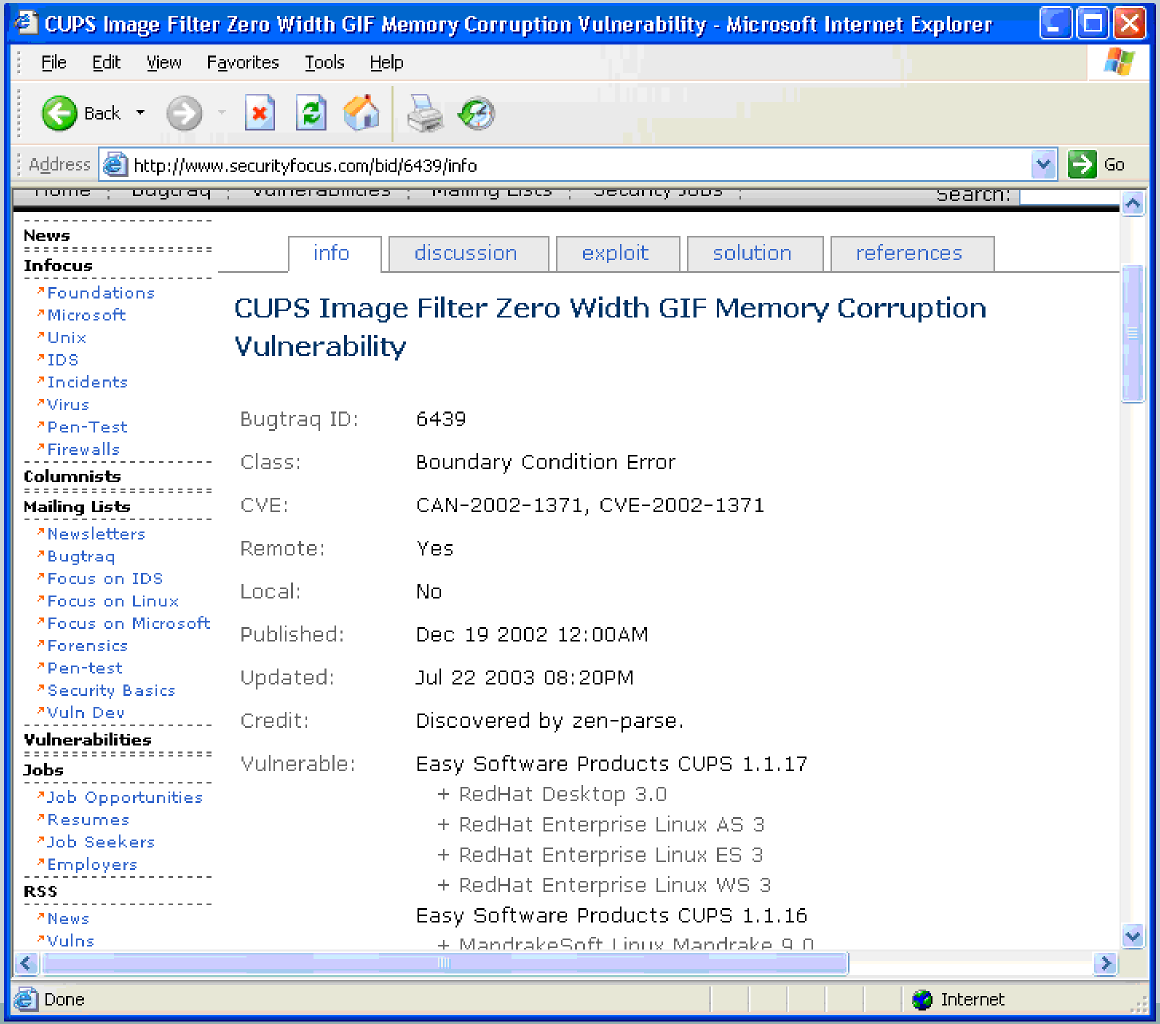
I. Description

The SNMP-to-DMI mapper daemon (snmpXdmi) translates Simple Network Management Protocol (SNMP) events to Desktop Management Interface (DMI) indications and vice-versa. Both protocols serve a similar purpose, and the translation daemon allows users to manage devices using either protocol. The snmpXdmi daemon registers itself with the snmpdx and dmid daemons, translating and forwarding requests from one daemon to the other. The snmpXdmi daemon, which is shipped with Solaris versions 2.6, 7 and 8, is enabled by default.

The snmpXdmi daemon contains a buffer overflow in the code for translating DMI indications to SNMP events. This buffer overflow is exploitable by remote intruders to gain root privileges.

More information about this vulnerability can be found in the advisory published by Job de Haas of ITSX:

Example of vulnerability publication by SF



CUPS Image Filter Zero Width GIF Memory Corruption Vulnerability

Bugtraq ID:	6439
Class:	Boundary Condition Error
CVE:	CAN-2002-1371, CVE-2002-1371
Remote:	Yes
Local:	No
Published:	Dec 19 2002 12:00AM
Updated:	Jul 22 2003 08:20PM
Credit:	Discovered by zen-parse.
Vulnerable:	Easy Software Products CUPS 1.1.17 + RedHat Desktop 3.0 + RedHat Enterprise Linux AS 3 + RedHat Enterprise Linux ES 3 + RedHat Enterprise Linux WS 3 Easy Software Products CUPS 1.1.16 + MandrakeSoft Linux Mandrake 9.0

Descriptive statistics

Patching Time (in days)	
Mean	56.5 (114.8)
Median	19
% patched	90%
Disclosure Time (in days) [*]	
Mean	15.6 (43.0)
% instant disclosure	67.1%
Number of observation facing disclosure	985

[*] Elapsed time between Vendor Notification and Disclosure (in days)

Descriptive statistics

Vendor characteristics (N=142)*

	Mean	Std. Dev.
Vendor employee size (in 000's)	17.60	66.12
Public firm	0.34	0.47
Open source	0.23	0.42

* Include only the vendors for which the reliable information is available

Vulnerability severity metric*

	Mean	Std. Dev.	Min	Max
Vulnerability severity metric	14.82	16.48	0	108.16
Number of affected vendors/vulnerability	8.23	21.53	1	242

* Vulnerability severity measurement by CERT/CC

Results

- Disclosure increases vendors' patching speed by 137%.
- Open source vendors patch 70% faster than closed source vendors.
- Severe vulnerabilities are patched faster.
- Vendors respond more favorably to CERT than SF. (almost 200% faster). In short vendors do not care as much if an individual or SF informs them about a vulnerability. Credibility of "who" is informing them matters.
- Small vendors do not patch as fast.

- Once we have these results, one can do a serious policy evaluation. For example, what is the optimal disclosure policy?

The Estimated Effect of Disclosure

Disclosure time *T*	Expected patching time (in days)		Effect of disclosure (in days)
	disclosed at time *T*	without disclosure	
0 (Instant disclosure)	33.03	61.77	28.74
1	36.90	61.77	24.87
2	38.60	61.77	23.17
3	39.98	61.77	21.79
4	41.16	61.77	20.61
5	42.21	61.77	19.56
6	43.18	61.77	18.59
7	44.07	61.77	17.70
8	44.90	61.77	16.87
9	45.69	61.77	16.08
10	46.43	61.77	15.34

These calculations are done setting the covariates at their average sample values namely that the vulnerable vendor is a public firm and has the average employee size of our sample; the vulnerability is published after 9/11 event, handled by CERT/CC and has the average severity metric of our sample.

EMPIRICAL Study (Romanosky, Telang, Acquisti 2010)

DATA BREACH NOTIFICATION LAWS

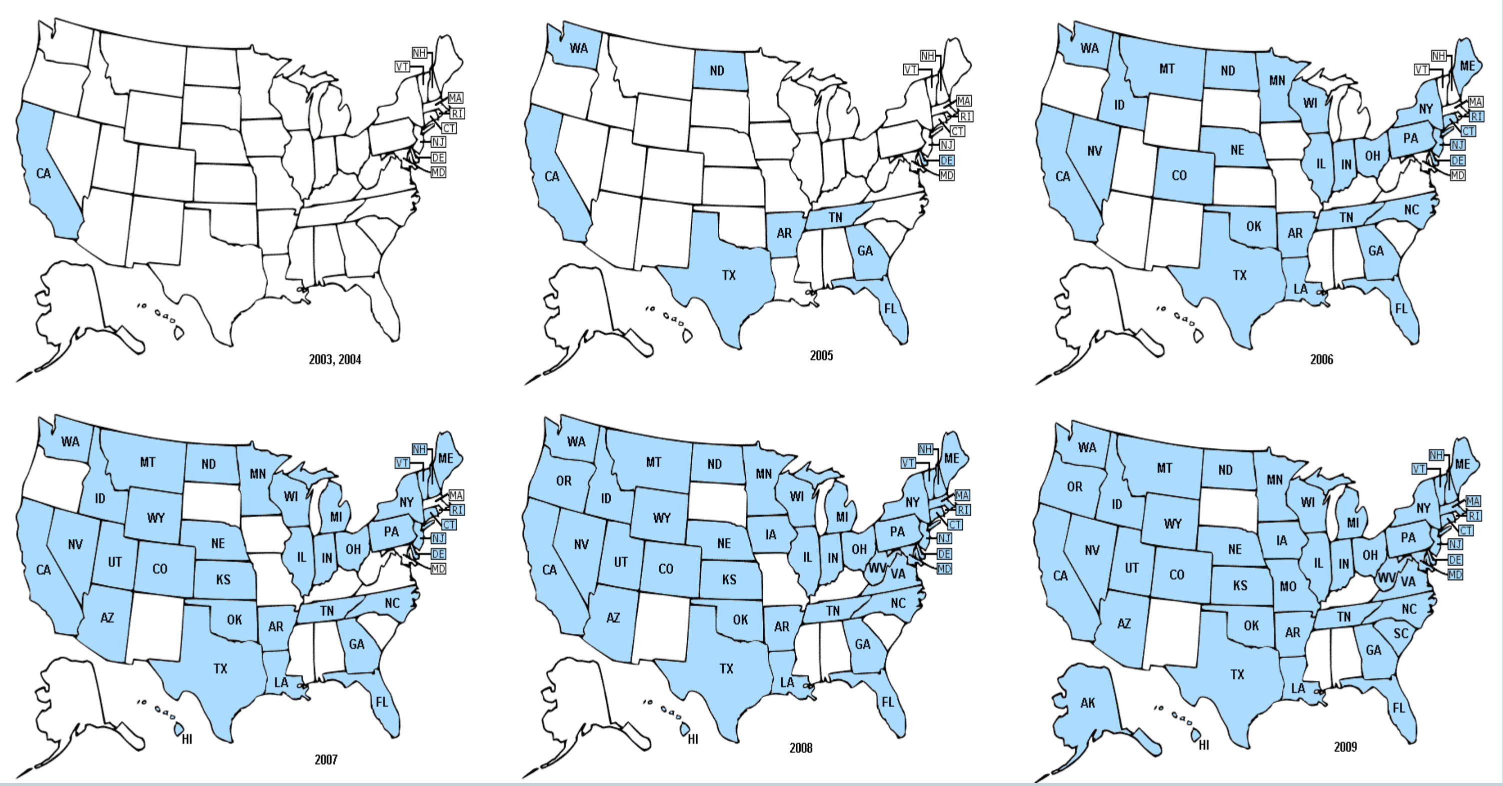
Data Breach Notification Laws

- Increasing number of data breaches have prompted many states to pass data breach disclosure laws.
  - One of the most important pieces of legislation in security and privacy space. Starting with CA in 2003, 45 states have adopted the law by 2009.
- Goals of the law:
  - Customer notification will allows them to take actions to prevent identity thefts. "to help consumers protect their financial security by requiring that state agencies and businesses [...] to quickly disclose to consumers any breach of the security of the, if the information disclosed could be used to commit identity theft" (SB1386).
  - Directly influence the incidences of identity thefts. "[t]he purpose of this Act is to alleviate the growing plague of identity theft...." (SB 2290).
- Again, the goal is that disclosure will force firms to do the right thing by investing in protecting customer data.

…not everyone agrees that disclosure laws are needed/effective

- Laws cause firms and consumers to incur unnecessary costs, leading to an overall worse outcome, esp. if the probability of idtheft is low (idAnalytics, 2006; Ponemon, 2008)
  - Many firms then have to replace the stolen card, set up 800 numbers to help customers, provide credit counseling etc.
  - A recent Ponemon report (2009) estimates these costs to be about $80 per stolen record.
- Consumers could become desensitized to numerous breach notifications, ignoring all of them (GAO, 2007)
- Stifles ecommerce and R&D by discouraging firms to innovate (Rubin and Lenard, 2005)

Adoption of state laws, 2002 - 2009



Identity theft data

- The FTC maintains a national database of consumer-reported identity theft complaints (1-877-ID-THEFT, www.ftc.gov)
- Uniform collection and management of data between states.
- Mined by law enforcement to catch offenders.
- Examples of idtheft :
  - Credit card charges (new, existing account, ~25%)
  - Loan, bank fraud (mortgage, car, etc, ~21%)
  - Phone and Utilities (unauthorized charges, new accounts, ~16%)
  - Government, medical benefits, etc (~10%)
- FTC only publishes data on annual basis. We invoked the Freedom of Information act to get monthly data which we aggregate semi-annually across 50 states.

Summary statistics

Variable (per 6-month period)	Mean	Std. Dev	Min	Max
Log(identity theft)	6.97	1.32	3.58	10.18
Identity theft rate (per 100,000)	32.00	13.49	5.67	84.74
Identity theft (total)	2,379.39	3,709.80	36	26,374
Has data breach law	0.38	0.48	0	1
Has FACTA	0.63	0.48	0	1
Has Credit Freeze Law	0.34	0.48	0	1
Per capita income	$35,547	$6,701	$23,019	$66,690
Unemployment rate	5.42	1.73	2.37	14.37
Log(population)	15.11	1.01	13.11	17.43
Newspaper articles	21.48	26.32	0	167
Log (fraud)	7.88	1.14	5.21	11.08

Results

Dep var: log(idtheft)	(1) Basic	(2) Basic + Controls	(3) Lagged	(4) Interstate	(5) Urban
Has Law	-0.050* (0.026)	0.061*** (0.023)		-0.047** (0.019)	-0.005 (0.028)
d1PerOld			-0.020 (0.015)		
d2PerOld			-0.037*** (0.012)		
d3PerOld			-0.023 (0.014)		
Has Law * Urban					-0.105*** (0.027)
Has FACTA		0.035* (0.019)	0.034* (0.018)	0.006 (0.011)	0.036* (0.019)
Has credit freeze law		0.036 (0.022)	0.020 (0.025)	0.032* (0.018)	0.039* (0.021)
Income per capita		-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Unemployment rate		0.003 (0.010)	0.002 (0.010)	0.006 (0.007)	0.008 (0.010)
Log (population)		-0.268 (0.343)	-0.300 (0.353)	-0.532* (0.278)	-0.092 (0.276)
State and time fixed effects	Y	Y	Y	Y	Y
Constant	6.852*** (0.014)	11.248** (5.317)	11.718** (5.490)	12.612*** (4.327)	8.359* (4.327)
Observations	800	800	800	800	800
R-squared	0.848	0.850	0.848	0.808	0.859
Number of states	50	50	50	50	50

Results

- Across all specifications, the estimate on Law is negative and significant. On average, the passage of law seems to have reduced the incidences of identity thefts by about 6%.
- Given the states and time fixed effects, most controls are not effective in predicting identity thefts.
- Some evidence of the lagged effect.
- Urban states have had a larger effect than the rural ones.
- Strictness of laws across different states does not seem to have an impact.

Significance of our results

- What is the economic significant of our estimate?
  - On average, the theft costs the victims $6383 (Javelin Research 2006).
  - A 6% reduction suggests that the marginal benefit of disclosure is about $380.
- Average cost of "notification" is about $30 per record (Javeline 2009). We are ignoring the ex-post costs (like consumer help etc.) which presumably reduce the cost to victims and hence a transfer.
- So the marginal cost of notification is $30.
- Every 13 notification should lead to one identity theft for the laws to be socially beneficial.
- Thus the theft probability per breached record should be about 7%.

To conclude…

- Disclosure has been and will continue to be an important policy tool in various contexts.
- However, empirical and experimental valuation of policy making is difficult due to lack of good data. Sometimes the data is available but not accessible to researchers.
- Without a thorough, rigorous and credible analysis, policy implementation and evaluation is always opinion driven as opposed to facts driven.